Correction du TD 9

Exercice 1 (Modèle Gamma-Poisson)

1. Notons π la densité de la loi a priori et $\pi(\cdot \,|\, \mathbf{X})$ la densité a posteriori. Pour tout $\theta \in \mathbb{R}_+^*$, on a

$$\pi(\theta \,|\, \mathbf{X}) \propto \theta^{p-1} e^{-\lambda\theta} \prod_{i=1}^{n} \frac{e^{-\theta}\theta^{X_i}}{X_i!} \propto \theta^{p+n\overline{X}_n-1} e^{-(n+\lambda)\theta} \mathbb{1}_{\theta>0} \,.$$

Ainsi la loi a posteriori $\Pi[\cdot \,|\, \mathbf{X}]$ est une loi Gamma$(p + n\overline{X}_n, n + \lambda)$.

2. Un estimateur de Bayes pour Π est donné par l'espérance a posteriori, soit ici

$$T^\star(\mathbf{X}) = \frac{p + n\overline{X}_n}{\lambda + n} \cdot$$

3. Pour tout $\theta > 0$, on peut effectuer une décomposition biais-variance, en tenant compte du fait que $\mathbb{E}_\theta[\overline{X}_n] = \theta$ et $\mathrm{Var}_\theta(\overline{X}_n) = \frac{\theta}{n}$, ce qui donne

$$\begin{aligned}\mathbf{R}(\theta, T^\star) &= \mathrm{Var}_\theta(T^\star) + (\mathbb{E}_\theta[T^\star] - \theta)^2 \\ &= \frac{n^2}{(n+\lambda)^2} \times \frac{\theta}{n} + \left(\frac{p+n\theta}{n+\lambda} - \theta\right)^2 \\ &= \frac{n\theta + \lambda^2\left(\theta - \frac{p}{\lambda}\right)^2}{(n+\lambda)^2} \,.\end{aligned}$$

4. Puisque $\boldsymbol{\theta} \sim$ Gamma(p, λ), on sait que $\mathbb{E}[\boldsymbol{\theta}] = p/\lambda$ et $\mathrm{Var}(\boldsymbol{\theta}) = p/\lambda^2$.
Première méthode
Puisque le risque ponctuel a été calculé à la question précédente, on peut obtenir le risque de Bayes pas intégration :

$$\mathbf{R}_B(\Pi) = \mathbf{R}_B(\Pi, T^\star) = \mathbb{E}\left[\mathbf{R}(\boldsymbol{\theta}, T^\star)\right] = \frac{1}{(n+\lambda)^2}\left(n\mathbb{E}[\boldsymbol{\theta}] + \lambda^2\mathbb{E}\left[\left(\boldsymbol{\theta} - \frac{p}{\lambda}\right)^2\right]\right)$$

donc

$$\mathbf{R}_B(\Pi) = \frac{1}{(n+\lambda)^2}\left(n\mathbb{E}[\boldsymbol{\theta}] + \lambda^2\,\mathrm{Var}(\boldsymbol{\theta})\right) = \frac{p}{\lambda(n+\lambda)}$$

Deuxième méthode
Puisque l'on considère la perte quadratique, le risque de Bayes est donné par

$$\mathbf{R}_B(\Pi) = \mathbf{R}_B(\Pi, T^\star) = \mathbb{E}[(m_{\mathbf{X}} - \boldsymbol{\theta})^2] = \mathbb{E}[\mathbb{E}[(m_{\mathbf{X}} - \boldsymbol{\theta})^2 \,|\, \mathbf{X}]] = \mathbb{E}[v_{\mathbf{X}}] = \mathbb{E}\left[\frac{p + n\overline{X}_n}{(n+\lambda)^2}\right].$$

En remarquant que $\mathbf{X} \,|\, \boldsymbol{\theta} = \theta$ est un n-échantillon de loi $\mathcal{P}(\theta)$, on obtient

$$\mathbb{E}[\overline{X}_n] = \mathbb{E}[\mathbb{E}[\overline{X}_n \,|\, \boldsymbol{\theta}]] = \mathbb{E}[\mathbb{E}[X_1 \,|\, \boldsymbol{\theta}]] = \mathbb{E}[\boldsymbol{\theta}] = \frac{p}{\lambda}.$$

Ainsi, $\mathbf{R}(\theta, T^\star) = \frac{p}{\lambda(n+\lambda)}$.

5. On cherche un minimiseur du risque a posteriori. Le risque a posteriori d'un estimateur T pour la perte ℓ s'écrit
$$\boldsymbol{\rho}(\Pi, T \mid \mathbf{X}) = \int_\Theta \theta^3 e^{-2\theta}(\theta - T)^2 \,\mathrm{d}\Pi(\theta \mid \mathbf{X})\,.$$
En utilisant que $\Pi[\cdot \mid \mathbf{X}] = \mathrm{Gamma}(p + n\overline{X}_n, \lambda + n)$, on obtient
$$\boldsymbol{\rho}(\Pi, T \mid \mathbf{X}) = \frac{(\lambda+n)^{p+n\overline{X}_n}}{\Gamma(p+n\overline{X}_n)} \int_0^{+\infty} (\theta - T)^2 \theta^{p+n\overline{X}_n+2} e^{-(\lambda+n+2)\theta} \,\mathrm{d}\theta\,.$$
Ainsi, la fonction $T \mapsto \boldsymbol{\rho}(\Pi, T \mid \mathbf{X})$ est proportionnelle à $\mathbb{E}\left[(\widetilde{\boldsymbol{\theta}} - T)^2 \mid \mathbf{X}\right]$ où la loi de $\widetilde{\boldsymbol{\theta}}$ sachant $\mathbf{X}$ est la loi $\mathrm{Gamma}(n\overline{X}_n + p + 3, \lambda + n + 2)$. Comme pour toute variable aléatoire Z de carré intégrable, la fonction $a \mapsto \mathbb{E}[(Z-a)^2]$ est minimale en $a = \mathbb{E}Z$, on obtient qu'un estimateur de Bayes pour Π et la perte ℓ est donné par
$$\widetilde{T}(\mathbf{X}) = \frac{n\overline{X}_n + p + 3}{\lambda + n + 2}\,.$$

Exercice 2 (Estimateurs de Bayes)

1. L'estimateur de Bayes pour la perte quadratique est la moyenne a posteriori $\widehat{\theta}(X) = \mathbb{E}[\boldsymbol{\theta} \mid X]$. Comme vu lors d'un TD précédent, un simple calcul montre que la loi a posteriori est
$$\Pi[\cdot \mid X] = \mathcal{N}\left(\frac{X}{\sigma^{-2}+1}, \frac{1}{\sigma^{-2}+1}\right)\,.$$
Ainsi $\widehat{\theta}(X) = \frac{X}{\sigma^{-2}+1}$ et, comme en exercice précédent,
$$\mathbf{R}_B(\Pi) = \mathbb{E}[v_X] = \frac{1}{\sigma^{-2}+1}.$$
Remarque : Tout comme ci-dessus, on peut commencer par calculer le risque ponctuel $R(\theta, \widehat{\theta}(X))$ puis écrire que $\mathbf{R}_B(\Pi) = \mathbb{E}\left[\mathbf{R}(\boldsymbol{\theta}, \widehat{\theta}(X))\right]$, ce qui correspond à intégrer ce risque ponctuel par rapport à la loi a priori de $\boldsymbol{\theta}$. Un calcul facile montre en fait que le risque ponctuel est constant : pour tout $\theta \in \mathbb{R}$,
$$\mathbf{R}(\theta, \widehat{\theta}(X)) = \frac{1}{\sigma^{-2}+1},$$
d'où le résultat.

2. Le raisonnement est le même que ci-dessus. Un estimateur de Bayes pour la perte ℓ peut s'obtenir en cherchant un minimiseur en T du risque a posteriori
$$\begin{aligned}\boldsymbol{\rho}(\Pi, T \mid X) = \int_\Theta \ell(\theta, T)\,\mathrm{d}\Pi(\theta \mid X) &\propto \int_{\mathbb{R}} (\theta - T)^2 \exp\left\{\frac{3\theta^2}{4} - \frac{\sigma^{-2}+1}{2}\left(\theta - \frac{X}{\sigma^{-2}+1}\right)^2\right\} \mathrm{d}\theta \\ &\propto \int_{\mathbb{R}} (\theta - T)^2 \exp\left\{-\frac{\sigma^{-2}-1/2}{2}\left(\theta - \frac{X}{\sigma^{-2}-1/2}\right)^2\right\} \mathrm{d}\theta\end{aligned}$$
L'intégrale est convergente si et seulement si $\sigma^{-2} > 1/2$. Dans ce cas, la quantité
$$\exp\left\{-\frac{\sigma^{-2}-1/2}{2}\left(\theta - \frac{X}{\sigma^{-2}-1/2}\right)^2\right\}$$

est le terme principal (non normalisé) de la densité d'une variable gaussienne $\mathcal{N}\left(\frac{X}{\sigma^{-2}-1/2}, \frac{1}{\sigma^{-2}-1/2}\right)$. Ainsi, la fonction $T \mapsto \boldsymbol{\rho}(\Pi, T \,|\, X)$ est proportionnelle à $\mathbb{E}\left[(\widetilde{\boldsymbol{\theta}} - T)^2 \,\middle|\, X\right]$ où la loi de $\widetilde{\boldsymbol{\theta}}$ sachant $\mathbf{X}$ est la loi $\mathcal{N}\left(\frac{X}{\sigma^{-2}-1/2}, \frac{1}{\sigma^{-2}-1/2}\right)$, donc

$$T^*(X) = \frac{X}{\sigma^{-2} - 1/2}$$

est un estimateur de Bayes pour la perte ℓ.
Pour calculer le risque de Bayes de T^*, on commence par calculer le risque ponctuel : pour tout réel θ, il s'écrit

$$\mathbf{R}(\theta, T^*) = \mathbb{E}_\theta[\ell(\theta, T^*)] = e^{\frac{3\theta^2}{4}} \mathbb{E}_\theta\left[(T^* - \theta)^2\right] = e^{\frac{3\theta^2}{4}} \left(\mathrm{Var}_\theta(T^*) + (\mathbb{E}_\theta[T^*] - \theta)^2\right).$$

Or on sait que $\mathbb{E}_\theta[X] = \theta$ et $\mathrm{Var}_\theta(X) = 1$, donc en posant $c := \sigma^{-2} - 1/2$, c'est-à dire $T^* = X/c$, on en déduit que

$$\mathbf{R}(\theta, T^*) = \frac{1 + (c-1)^2\theta^2}{c^2} e^{\frac{3\theta^2}{4}}.$$

Il s'ensuit que

$$\mathbf{R}_B(\Pi, T^*) = \mathbb{E}\left[\mathbf{R}(\boldsymbol{\theta}, T^*)\right] = \frac{1}{c^2}\mathbb{E}\left[(1 + (c-1)^2\boldsymbol{\theta}^2)e^{\frac{3\boldsymbol{\theta}^2}{4}}\right] = \frac{1}{c^2}\int_{\mathbb{R}} (1+(c-1)^2\theta^2)e^{\frac{3\theta^2}{4}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\theta^2}{2\sigma^2}} \,\mathrm{d}\theta.$$

Ceci montre que le risque de Bayes est infini si $\sigma^2 \geq 2/3$, puisque dans ce cas cette intégrale est divergente. Si $\sigma^2 < 2/3$, on fait à nouveau apparaître une loi gaussienne en notant que

$$e^{\frac{3\theta^2}{4}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\theta^2}{2\sigma^2}} = \frac{1}{\sqrt{1 - 3\sigma^2/2}} \sqrt{\frac{\sigma^{-2} - 3/2}{2\pi}} e^{-\frac{\sigma^{-2}-3/2}{2}\theta^2}.$$

Soit $Y \sim \mathcal{N}(0, 1/(\sigma^{-2} - 3/2))$, alors $\mathbb{E}[Y^2] = 1/(\sigma^{-2} - 3/2) = 1/(c-1)$ donc

$$\mathbf{R}_B(\Pi, T^*) = \frac{1}{c^2\sqrt{1 - 3\sigma^2/2}}\mathbb{E}\left[(1 + (c-1)^2Y^2)\right] = \frac{1}{c\sqrt{1 - 3\sigma^2/2}} = \frac{1}{(\sigma^{-2} - 1/2)\sqrt{1 - 3\sigma^2/2}}.$$

Exercice 3 (De Bayes et de risque constant implique minimax)

1. La formule de Bayes donne, pour tous $a, b > 0$,

$$\pi_{a,b}(\theta \,|\, \mathbf{X}) \propto \theta^{a-1}(1-\theta)^{b-1}\mathbb{1}_{0<\theta<1}\theta^{n\overline{X}_n}(1-\theta)^{n-n\overline{X}_n} \propto \theta^{a+n\overline{X}_n-1}(1-\theta)^{b+n-n\overline{X}_n-1}\mathbb{1}_{0<\theta<1}.$$

La loi a posteriori est une loi $\mathrm{Beta}(a + n\overline{X}_n, b + n - n\overline{X}_n)$. Par le cours, un estimateur de Bayes $\widehat{\theta}_{a,b}(\mathbf{X})$ pour $\Pi_{a,b}$ et la perte quadratique est la moyenne a posteriori. En utilisant que l'espérance d'une loi $\mathrm{Beta}(a, b)$ est égale à $\frac{a}{a+b}$, on obtient

$$\widehat{\theta}_{a,b}(\mathbf{X}) = \frac{a + n\overline{X}_n}{a + b + n}.$$

2. D'après le cours, un estimateur de Bayes de risque constant est minimax. Il suffit donc de trouver, parmi les estimateurs $\left(\widehat{\theta}_{a,b}(\mathbf{X})\right)_{a,b>0}$ (qui sont de Bayes), un estimateur qui soit de risque constant.

L'énoncé suggère de poser $a = b$, on calcule donc le risque quadratique de l'estimateur $\widehat{\theta}_{a,a}(\mathbf{X})$,

$$\begin{aligned}
\mathbb{E}_\theta\left[(\widehat{\theta}_{a,a}(\mathbf{X}) - \theta)^2\right] &= \mathbb{E}_\theta\left[\left(\frac{n(\overline{X}_n - \theta)}{2a+n} - \frac{2a\theta - a}{2a+n}\right)^2\right] \\
&= \frac{n^2 \operatorname{Var}_\theta(\overline{X}_n)}{(2a+n)^2} + \frac{a^2(2\theta-1)^2}{(2a+n)^2} \\
&= (2a+n)^{-2}\left(n\theta(1-\theta) + a^2(2\theta-1)^2\right) \\
&= (2a+n)^{-2}\left((4a^2-n)\theta^2 + (n-4a^2)\theta + a^2\right).
\end{aligned}$$

On en conclut que si l'on choisit a tel que $4a^2 = n$, soit

$$a_n = \frac{\sqrt{n}}{2},$$

l'estimateur

$$\widehat{\theta}_{a_n,a_n}(\mathbf{X}) = \frac{\frac{1}{2} + \sqrt{n}\overline{X}_n}{1+\sqrt{n}}$$

est de risque constant et de Bayes donc, d'après un théorème du cours, il est minimax.

3. Pour savoir si T est minimax, il suffit de comparer $\mathbf{R}_{\max}(T)$ au risque minimax $\mathbf{R}_M$. Or on peut calculer ce dernier car la question précédente a fourni un estimateur minimax, à savoir $\widehat{\theta}_{a_n,a_n}(\mathbf{X})$. Le risque maximal de ce dernier est, d'après la question précédente, puisqu'il est de risque constant, égal à

$$\mathbf{R}_{\max}(\widehat{\theta}_{a_n,a_n}(\mathbf{X})) = \frac{a_n^2}{(2a_n+n)^2} = \frac{n}{4(n+\sqrt{n})^2}.$$

Or, le risque maximal de T vaut

$$\mathbf{R}_{\max}(T) = \sup_{\theta\in[0,1]} \mathbf{R}(\theta, T) = \sup_{\theta\in[0,1]} \operatorname{Var}_\theta(\overline{X}_n) = \sup_{\theta\in[0,1]} \frac{\theta(1-\theta)}{n} = \frac{1}{4n},$$

On en conclut $\mathbf{R}_{\max}(T) > \mathbf{R}_{\max}(\widehat{\theta}_{a_n,a_n}(\mathbf{X})) = \mathbf{R}_M$, donc T n'est pas minimax.

Exercice 4 (Contrôle continu 2020)

1. Si $Y \sim \mathcal{N}(0, s^2)$, alors par parité de la fonction intégrée

$$\mathbb{E}[|Y|] = 2\int_0^\infty y\frac{1}{\sqrt{2\pi s^2}}e^{-\frac{y^2}{2s^2}}dy = \sqrt{\frac{2}{\pi}}s\left[-e^{-\frac{y^2}{2s^2}}\right]_0^\infty = \sqrt{\frac{2}{\pi}}s.$$

2. Un calcul classique donne $\Pi[\cdot|\mathbf{X}] = \mathcal{N}\left(\frac{\mu\sigma^{-2}+n\overline{X}_n}{n+\sigma^{-2}}, \frac{1}{n+\sigma^{-2}}\right)$.
3. Par un résultat du cours, un estimateur de Bayes pour la perte en valeur absolue est donné par la médiane a posteriori de $\boldsymbol{\theta}$, qui correspond ici à sa moyenne, i.e. $T^*(\mathbf{X}) = \frac{\mu\sigma^{-2}+n\overline{X}_n}{n+\sigma^{-2}}$.
4. On a, en utilisant la question préliminaire,

$$\mathbf{R}_B(\Pi, T^*) = \mathbb{E}\left[|\boldsymbol{\theta} - \mathbb{E}[\boldsymbol{\theta}|\mathbf{X}]|\right] = \mathbb{E}\left[\mathbb{E}\left[|\boldsymbol{\theta} - \mathbb{E}[\boldsymbol{\theta}|\mathbf{X}]|\,|\mathbf{X}\right]\right] = \mathbb{E}\left[\left|\mathcal{N}\left(0, \frac{1}{n+\sigma^{-2}}\right)\right|\right] = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{\sigma^{-2}+n}}.$$

5. Le risque ponctuel de l'estimateur $\overline{X}_n$ est donné par

$$\mathbf{R}(\theta, \overline{X}_n) = \mathbb{E}_\theta\left[|\overline{X}_n - \theta|\right] = \frac{1}{\sqrt{n}}\mathbb{E}\left[|\mathcal{N}(0,1)|\right] = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{n}}.$$

Ainsi, le risque de $\overline{X}_n$ est constant donc $\mathbf{R}_{\max}(\overline{X}_n) = \sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{n}}$. En considérant la suite de lois a priori $(\Pi_k)_{k\geq 1}$ avec $\Pi_k = \mathcal{N}(0,k)$, on obtient

$$\mathbf{R}_{\max}(\overline{X}_n) = \lim_{k\to+\infty}\sqrt{\frac{2}{\pi}}\frac{1}{\sqrt{\frac{1}{k}+n}} = \lim_{k\to+\infty}\mathbf{R}_B(\Pi_k)\,.$$

Par une Proposition du cours, on en déduit que $\overline{X}_n$ est minimax pour la perte en valeur absolue.

Exercice 5 (Contrôle continu 2020)

1. On a

$$\begin{aligned}\pi(\theta|\mathbf{X}) &\propto \theta^{a-1}(1-\theta)^{b-1}\mathbf{1}_{[0,1]}(\theta)\prod_{i=1}^{n}(X_i+1)(1-\theta)^2\theta^{X_i}\\ &\propto \theta^{a+n\overline{X}_n-1}(1-\theta)^{b+2n-1}\mathbf{1}_{[0,1]}(\theta)\end{aligned}$$

et on a donc $\boldsymbol{\theta}|\mathbf{X} \sim \text{Beta}\left(a+n\overline{X}_n, b+2n\right)$. Ainsi,

$$\mathbb{E}\left[\boldsymbol{\theta}|\mathbf{X}\right] = \frac{a+n\overline{X}_n}{a+b+2n+n\overline{X}_n}, \quad \mathbb{V}\left(\boldsymbol{\theta}|\mathbf{X}\right) = \frac{(a+n\overline{X}_n)(b+2n)}{\left(a+b+2n+n\overline{X}_n\right)^2\left(a+b+2n+n\overline{X}_n+1\right)}.$$

2. Avec les calculs usuels on obtient l'intervalle de crédibilité

$$IC_{1-\alpha}(\mathbf{X}) = \left[\mathbb{E}\left[\boldsymbol{\theta}|\mathbf{X}\right] \pm \sqrt{\frac{\mathbb{V}\left(\boldsymbol{\theta}|\mathbf{X}\right)}{\alpha}}\right]$$

3. Un estimateur de Bayes pour la perte quadratique est donné par la moyenne a posteriori $\mathbb{E}[\boldsymbol{\theta}|\mathbf{X}]$.
4. On considère maintenant la perte $\ell(\theta,T) = \frac{1}{(1-\theta)^2}(\theta-T)^2$.
   (a) On cherche un élément T^* qui minimise le risque a posteriori, i.e. la fonction $T \mapsto \boldsymbol{\rho}(\Pi, T \mid X) = \mathbb{E}[\ell(\boldsymbol{\theta},T)|\mathbf{X}]$. On a

$$\begin{aligned}\int_\Theta \ell(\theta,T)d\Pi(\theta|\mathbf{X}) &\propto \int_0^1 \frac{(\theta-T)^2}{(1-\theta)^2}\theta^{a+n\overline{X}_n-1}(1-\theta)^{b+2n-1}d\theta\\ &\propto \int_0^1 (\theta-T)^2\theta^{a+n\overline{X}_n-1}(1-\theta)^{b+2n-3}d\theta\end{aligned}$$

   et comme $n\geq 1$, $\theta^{a+n\overline{X}_n-1}(1-\theta)^{b+2n-3}\mathbf{1}_{[0,1]}(\theta)$ est le terme général de la densité d'une loi $\text{Beta}(a+n\overline{X}_n, b+2(n-1))$, et on obtient donc l'estimateur de Bayes

$$T^*(\mathbf{X}) = \frac{a+n\overline{X}_n}{a+b+n\overline{X}_n+2(n-1)}.$$

(b) On souhaite estimer

$$\mathbf{I} = \mathbf{R}_B(\Pi) = \mathbb{E}\left[\ell\left(\boldsymbol{\theta}, T^*(X_1, \ldots, X_n)\right)\right] .$$

Pour cela, on prend $N \in \mathbb{N}^*$, et on simule N vecteurs $(\boldsymbol{\theta}_j, X_1^j, \ldots, X_n^j)$, $j = 1, \ldots, N$, i.i.d. de même loi que $(\boldsymbol{\theta}, X_1, \ldots, X_n)$. Cela est possible en simulant d'abord $\boldsymbol{\theta}_j$ de loi Beta(a, b). Puis, sachant $\boldsymbol{\theta}_j$, en simulant $X_1^j, \ldots, X_n^j$ i.i.d. de loi $P_{\boldsymbol{\theta}_j}$ (loi binomiale négative de paramètres 2 et $1 - \boldsymbol{\theta}_j$). Ensuite, pour $j = 1, \ldots, N$, on calcule

$$T_j^* = T_j^*(X_1^j, \ldots, X_n^j) = \frac{a + \sum_{i=1}^n X_i^j}{a + b + 2(n-1) + \sum_{i=1}^n X_i^j} ,$$

et

$$\ell(\boldsymbol{\theta}_j, T_j^*) = \frac{1}{(1 - \boldsymbol{\theta}_j)^2}(\boldsymbol{\theta}_j - T_j^*)^2 .$$

On approche alors $\mathbf{I}$ par l'estimateur Monte-Carlo classique

$$\widehat{\mathbf{I}}_N = \frac{1}{N} \sum_{j=1}^{N} \ell(\boldsymbol{\theta}_j, T_j^*) ,$$

qui est fortement consistant par la loi forte des grands nombres.

Remarque : Ce dernier point suppose que $\mathbf{I} = \mathbf{R}_B(\Pi) < \infty$, sans quoi l'estimateur Monte-Carlo tend presque sûrement vers l'infini. Ceci découle du fait que T^* minimise la fonction $T \mapsto \mathbb{E}[\ell(\boldsymbol{\theta}, T)|\mathbf{X}]$. En particulier, il fait mieux que l'estimateur trivial $\check{T} \equiv 1$, pour lequel $\mathbb{E}[\ell(\boldsymbol{\theta}, \check{T})|\mathbf{X}] \equiv 1$ et dont le risque de Bayes, qui majore celui de T^*, vaut donc 1.

Exercice 6 (Rattrapage 2020)

1. Pour tout $\theta > 0$, on a

$$\begin{aligned}\pi_{a,b}(\theta|\mathbf{X}) &\propto \theta^{a-1} e^{-b\theta} \mathbb{1}_{\theta>0} \prod_{i=1}^{n} \left(\theta e^{-\theta X_i}\right) \\ &\propto \theta^{n+a-1} e^{-(b+n\overline{X}_n)\theta} \mathbb{1}_{\theta>0} .\end{aligned}$$

Ainsi $\Pi_{a,b}[\cdot \,|\, \mathbf{X}] = \Gamma\left(a+n, b+n\overline{X}_n\right)$. En particulier,

$$m_{\mathbf{X}} = \frac{a+n}{b+n\overline{X}_n} \qquad \text{et} \qquad v_{\mathbf{X}} = \frac{a+n}{(b+n\overline{X}_n)^2} .$$

2. On procède par minimisation du risque a posteriori. Pour tout estimateur T, on a

$$\boldsymbol{\rho}(\Pi, T \,|\, \mathbf{X}) = \mathbb{E}\left[\ell(\boldsymbol{\theta}, T(\mathbf{X})) \,\middle|\, \mathbf{X}\right] = \int_0^{+\infty} \frac{(b+n\overline{X}_n)^{a+n}}{\Gamma(a+n)} (T(\mathbf{X}) - \theta)^2 \theta^{a+n-3} e^{-(b+n\overline{X}_n)\theta} \, \mathrm{d}\theta .$$

À constante près, on reconnaît $\int_0^{+\infty} (T(\mathbf{X}) - \theta)^2 \, \mathrm{d}\widetilde{\Pi}_{\mathbf{X}}(\theta)$ avec $\widetilde{\Pi}_{\mathbf{X}} = \Gamma(a+n-2, b+n\overline{X}_n)$. On en déduit qu'un estimateur de Bayes est donné par

$$T_{a,b}(\mathbf{X}) = \frac{a+n-2}{b+n\overline{X}_n} ,$$

qui correspond à l'espérance (conditionnelle à $\mathbf{X}$) de la loi $\widetilde{\Pi}_{\mathbf{X}}$.

3. Au vu de la question précédente, on peut chercher à exprimer le risque a posteriori de $T_{a,b}$ en fonction de la variance de la loi $\widetilde{\Pi}_{\mathbf{X}}$. On a

$$\begin{aligned}
\boldsymbol{\rho}(\Pi, T_{a,b}(\mathbf{X}) \,|\, \mathbf{X}) &= \frac{(b+n\overline{X}_n)^{a+n}}{\Gamma(a+n)} \int_0^{+\infty} (T_{a,b}(\mathbf{X})-\theta)^2 \theta^{a+n-3} e^{-(b+n\overline{X}_n)\theta} \,\mathrm{d}\theta \\
&= \frac{(b+n\overline{X}_n)^2}{(a+n-1)(a+n-2)} \int_0^{+\infty} \frac{(b+n\overline{X}_n)^{a+n-2}}{\Gamma(a+n-2)} (T_{a,b}(\mathbf{X})-\theta)^2 \theta^{a+n-3} e^{-(b+n\overline{X}_n)\theta} \,\mathrm{d}\theta \\
&= \frac{(b+n\overline{X}_n)^2}{(a+n-1)(a+n-2)} \frac{a+n-2}{(b+n\overline{X}_n)^2} \\
&= \frac{1}{a+n-1} \,.
\end{aligned}$$

   Ainsi le risque de Bayes vaut $\mathbf{R}_B(\Pi_{a,b}) = \frac{1}{a+n-1}$.

4. Soit $\theta > 0$. On commence par calculer le risque ponctuel. En utilisant l'indication de l'énoncé, on a

$$\begin{aligned}
\mathbf{R}(\theta, T^*) &= \mathbb{E}_\theta\left[\left(\frac{n-2}{\theta\sum_{i=1}^n X_i} - 1\right)^2\right] \\
&= \mathbb{E}_\theta\left[\left(\frac{n-2}{Z} - 1\right)^2\right] ,
\end{aligned}$$

   où $Z \sim \Gamma(n,1)$. Par le rappel de l'énoncé, on a

$$\mathbb{E}\left[\frac{1}{Z}\right] = \frac{\Gamma(n-1)}{\Gamma(n)} = \frac{1}{n-1} \,.$$

   et

$$\mathbb{E}\left[\frac{1}{Z^2}\right] = \frac{\Gamma(n-2)}{\Gamma(n)} = \frac{1}{(n-1)(n-2)} \,.$$

   Ainsi,

$$\begin{aligned}
\mathbf{R}(\theta, T^*) &= (n-2)^2 \mathbb{E}\left[\frac{1}{Z^2}\right] - 2(n-2)\mathbb{E}\left[\frac{1}{Z}\right] + 1 \\
&= (n-2)^2 \frac{1}{(n-1)(n-2)} - 2(n-2)\frac{1}{n-1} + 1 \\
&= \frac{1}{n-1} \,.
\end{aligned}$$

   Le risque de T^* est donc constant égal à $\frac{1}{n-1}$. En particulier, $\mathbf{R}_{\max}(T^*) = \frac{1}{n-1}$.

5. En considérant la suite de lois a priori $(\Pi_{1/k,1})_{k\geq 1}$, on a

$$\lim_{k\to+\infty} \mathbf{R}_B(\Pi_{1/k,1}) = \lim_{k\to+\infty} \frac{1}{1/k+n-1} = \frac{1}{n-1} = \mathbf{R}_{\max}(T^*) \,.$$

   Par une proposition du cours, on en déduit que T^* est minimax.

Exercice 7 (Quantiles)

On peut commencer par remarquer que dans le cas où $a = b = 1$, il s'agit de la fonction de perte en valeur absolue, pour laquelle nous avons vu en cours que l'estimateur de Bayes est la médiane a posteriori.

Pour trouver un estimateur de Bayes pour Π, il suffit de trouver un minimiseur du risque a posteriori soit

$$T^*(X) = \arg\min_t \int_\Theta \ell(\theta,t)\,\mathrm{d}\Pi(\theta \mid X) = \arg\min_t \mathbb{E}\left[\ell(\boldsymbol{\theta},t) \,\middle|\, X\right].$$

Par hypothèse, la loi a priori (donc aussi la loi a posteriori) a une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$. En particulier, la fonction de répartition $F = F_{\boldsymbol{\theta}|X}$ de la loi a posteriori est continue. Par définition de ℓ, on a, pour tout $t \in \mathbb{R}$,

$$\mathbb{E}\left[\ell(\boldsymbol{\theta},t) \,\middle|\, X\right] = \mathbb{E}\left[b(\boldsymbol{\theta}-t)\mathbb{1}_{\boldsymbol{\theta}>t} \,\middle|\, X\right] + \mathbb{E}\left[a(t-\boldsymbol{\theta})\mathbb{1}_{\boldsymbol{\theta}\le t} \,\middle|\, X\right].$$

Notons que, $\boldsymbol{\theta}$ étant par hypothèse intégrable pour la loi a priori Π, les deux termes sont finis presque sûrement car positifs et d'espérance finie. En effet, on peut par exemple écrire pour le premier :

$$\mathbb{E}\left[\mathbb{E}\left[b(\boldsymbol{\theta}-t)\mathbb{1}_{\boldsymbol{\theta}>t} \,\middle|\, X\right]\right] = \mathbb{E}\left[b(\boldsymbol{\theta}-t)\mathbb{1}_{\boldsymbol{\theta}>t}\right] \le b\mathbb{E}[|\boldsymbol{\theta}-t|] \le b(\mathbb{E}[|\boldsymbol{\theta}|]+|t|) < \infty.$$

Ensuite, pour toute variable aléatoire positive U, on a $\mathbb{E}[U] = \int_0^\infty \mathbb{P}(U>u)\,\mathrm{d}u$, donc

$$\begin{aligned}
\mathbb{E}\left[\ell(\boldsymbol{\theta},t) \,\middle|\, X\right] &= \int_0^{+\infty} \mathbb{P}\left(b(\boldsymbol{\theta}-t) > u \,\middle|\, X\right)\,\mathrm{d}u + \int_0^{+\infty} \mathbb{P}\left(a(t-\boldsymbol{\theta}) > u \,\middle|\, X\right)\,\mathrm{d}u \\
&= \int_0^{+\infty} \left(1 - F\left(t+\frac{u}{b}\right)\right)\mathrm{d}u + \int_0^{+\infty} F\left(t-\frac{u}{a}\right)\mathrm{d}u \\
&= b\int_t^{+\infty} (1-F(s))\,\mathrm{d}s + a\int_{-\infty}^t F(s)\,\mathrm{d}s\,,
\end{aligned}$$

où l'on a effectué les changements de variables $s = t + \frac{u}{b}$ et $s = t - \frac{u}{a}$ respectivement. Rappelons que les deux intégrales sont finies pour tout réel t. Etudions la seconde : on peut donc écrire

$$\int_{-\infty}^t F(s)\,\mathrm{d}s = \int_{-\infty}^0 F(s)\,\mathrm{d}s + \int_0^t F(s)\,\mathrm{d}s = C + \int_0^t F(s)\,\mathrm{d}s.$$

Comme F est continue, elle admet une primitive Φ, qui est donc de classe C^1 et vérifie

$$\int_{-\infty}^t F(s)\,\mathrm{d}s = C + \Phi(t) - \Phi(0).$$

Par conséquent, la seconde intégrale, vue comme une fonction de t, est dérivable sur $\mathbb{R}$, de dérivée

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(a\int_{-\infty}^t F(s)\,\mathrm{d}s\right) = a\Phi'(t) = aF(t).$$

Un raisonnement similaire donne

$$\int_t^{+\infty}(1-F(s))\,\mathrm{d}s = \int_0^{+\infty}(1-F(s))\,\mathrm{d}s - \int_0^t (1-F(s))\,\mathrm{d}s = C' - t + \Phi(t) - \Phi(0),$$

si bien que

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(b\int_t^{+\infty}(1-F(s))\,\mathrm{d}s\right) = b(F(t)-1).$$

Au total, la fonction $t \mapsto \mathbb{E}\left[\ell(\boldsymbol{\theta},t) \,\middle|\, X\right]$ est C^1 sur $\mathbb{R}$, de dérivée

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\ell(\boldsymbol{\theta},t) \,\middle|\, X\right] = (a+b)F(t) - b.$$

Pour trouver son minimum, on écrit

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}\left[\ell(\boldsymbol{\theta},t)\,\middle|\,X\right] \geq 0 \;\Leftrightarrow\; F(t) \geq \frac{b}{a+b}.$$

Notons comme d'habitude F^{-1} l'inverse généralisée (ou fonction quantile) de F. On rappelle que pour tout $0 < u < 1$, le quantile d'ordre u est défini par

$$F^{-1}(u) = \inf\{t \in \mathbb{R},\ F(t) \geq u\},$$

et que l'on a l'équivalence suivante :

$$F(t) \geq u \;\Leftrightarrow\; t \geq F^{-1}(u).$$

En particulier, $t < F^{-1}(u)$ implique $F(t) < u$. Rappelons aussi que, si F est continue (comme c'est le cas ici), on sait que $F(F^{-1}(u)) = u$ pour tout $0 < u < 1$. Pour ce qui nous concerne, en notant $T^*(X)$ le quantile d'ordre $b/(a+b)$ de F, il s'ensuit que la dérivée de $t \mapsto \mathbb{E}\left[\ell(\boldsymbol{\theta},t)\,\middle|\,X\right]$ est strictement négative si $t < T^*(X)$ et positive (au sens large) si $t \geq T^*(X)$. Ce dernier est donc un minimiseur de la fonction $t \mapsto \mathbb{E}\left[\ell(\boldsymbol{\theta},t)\,\middle|\,X\right]$. Au total, $T^*(X)$ est un estimateur de Bayes pour la fonction de perte ℓ. Ceci généralise bien le cas $a = b = 1$ vu en cours.

Remarque : La fonction F est continue mais rien n'empêche qu'elle soit constante sur un intervalle de la forme $[T^*(X), T^*(X)+h]$, auquel cas toute valeur $\widetilde{T}(X) \in [T^*(X), T^*(X)+h]$ est aussi un estimateur de Bayes. Il n'y a donc pas nécessairement unicité.